52

53

(2) INFORMATION FOR SEQ ID NO:1:

Raw Sequence Listing (1814) 04/15/92 Patent Application US/07/659,408A

SEQUENCE LISTING

#### 1 2 3 (1) GENERAL INFORMATION: 6 (i) APPLICANT: CAPUT, DANIEL 7 FERRARA, PASCUAL 8 GUILLEMOT, JEAN-CLAUDE 9 KAGHAD, MOURAD 10 LEGOUX, RICHARD 11 LOISON, GERARD 12 LARBRE, ELIZABETH 13 LUPKER, JOHANNES 14 LEPLATOIS, PASCUAL 15 SALOME, MARK 16 17 (ii) TITLE OF INVENTION: URATE OXIDASE ACTIVITY PROTEIN, 18 RECOMBINANT GENE CODING THEREFOR, EXPRESSION VECTOR, 19 MICRO-ORGANISMS AND TRANSFORMED CELLS 20 21 (iii) NUMBER OF SEQUENCES: 36 22 23 (iv) CORRESPONDENCE ADDRESS: 24 (A) ADDRESSEE: Foley & Lardner 25 (B) STREET: 1800 Diagonal Road, Suite 500 26 (C) CITY: Alexandria 27 (D) STATE: Virginia (E) COUNTRY: USA 28 29 (F) ZIP: 22313-0299 30 31 (v) COMPUTER READABLE FORM: 32 (A) MEDIUM TYPE: Floppy disk 33 (B) COMPUTER: IBM PC compatible 34 (C) OPERATING SYSTEM: PC-DOS/MS-DOS 35 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 36 37 (vi) CURRENT APPLICATION DATA: 38 (A) APPLICATION NUMBER: US 07/659,408 39 (B) FILING DATE: 25-APR-1991 40 (C) CLASSIFICATION: 41 42 (viii) ATTORNEY/AGENT INFORMATION: 43 (A) NAME: BENT, Stephen A. 44 (B) REGISTRATION NUMBER: 29,768 45 (C) REFERENCE/DOCKET NUMBER: 16781/276 BEDL 46 47 (ix) TELECOMMUNICATION INFORMATION: 48 (A) TELEPHONE: (703)836-9300 49 (B) TELEFAX: (703)683-4109 50 (C) TELEX: 899149 51

#### Raw Sequence Listing

04/15/92 09:05:35

34																
55 56 57	(i)	(Ā	) LE		: 30	1 am.	STIC:		8							
58		•	•	POLO												
59		(1)	, 10.	CLO	<b>JI.</b>	TTHE	ar									
60 61	(ii)	MOL	ECUL	E TY	PE: 1	prot	ein									
62 63	(iii)	HYP	OTHE:	TICA	L: No	0										
64 65 66	(vi)						rgil:	lus :	flav	us						
67 68 69 70	(vii)						xida	se								
71 72	(xi)	SEQ	UENC	E DE	SCRI	PTIO	N: S1	EQ II	ои о	:1:						
73	Ser	Ala	Val	Lys	Ala	Ala	Arg	Tyr	Gly	Lys	Asp	Asn	Val	Arg	Val	Tyr
74	1				5					10					15	
75	_			_							_	_	_			
76 77 78	Lys	Val	His	Lys 20	Asp	Glu	Lys	Thr	Gly 25	Val	Gln	Thr	Val	Tyr 30	Glu	Met
79 80 81	Thr	Val	Сув 35	Val	Leu	Leu	Glu	Gly 40	Glu	Ile	Glu	Thr	Ser 45	Tyr	Thr	Lys
82 83 84	Ala	Asp 50	Asn	Ser	Val	Ile	Val 55	Ala	Thr	Asp	Ser	Ile 60	Lys	Asn	Thr	Ile
85 86 87	Tyr 65	Ile	Thr	Ala	Lys	Gln 70	Asn	Pro	Val	Thr	Pro 75	Pro	Glu	Leu	Phe	Gly 80
88 89 90	Ser	Ile	Leu	Gly	Thr 85	His	Phe	Ile	Glu	Dys 90	Tyr	Asn	His	Ile	His 95	Ala
91 92 93	Ala	His	Val	Asn 100	Ile	Val	Сув	His	Arg 105	Trp	Thr	Arg	Met	Asp 110	Ile	Asp
94 95 96	Gly	Lys	Pro 115	His	Pro	His	Ser	Phe 120	Ile	Arg	Asp	Ser	Glu 125	Glu	Lys	Arg
97 98 99	Asn	Val 130	Gln	Val	Asp	Val	Val 135	Glu	Gly	Lys	Gly	Ile 140	Asp	Ile	Lys	Ser
100 101 102	Ser 145	Leu	Ser	Gly	Leu	Thr 150	Val	Leu	Lys	Ser	Thr 155	Asn	Ser	Gln	Phe	Trp 160
103 104 105	Gly	Phe	Leu	Arg	Asp 165	Glu	Tyr	Thr	Thr	Leu 170	Lys	Glu	Thr	Trp	Asp 175	Arg
106	Ile	Leu	Ser	Thr	Asp	Val	Asp	Ala	Thr	Trp	Gln	Trp	Lys	Asn	Phe	Ser

#### Raw Sequence Listing

04/15/92 09:05:37

107				180					185					190		
108																
109	Gly	Leu		Glu	Val	Arg	Ser		Val	Pro	Lys	Phe	-	Ala	Thr	Trp
110			195					200					205			
111	21-	m	71-	N	<b>01</b>	**- 1	mъ	<b>T</b>	*	mh	nh -		<b>01</b>	<b>3</b>	<b>3</b>	
112 113	ATA	210	ATA	Arg	GIU	vaı	215	Leu	rys	Thr	Pne	220	GIU	Asp	Asn	ser
114		210					215					220				
115	Ala	Ser	۷a۱	Gln	Ala	Thr	Met	Tvr	Lvs	Met	Ala	Glu	Gln	Ile	Leu	Ala
116	225		,	<b></b>	****	230		-1-	2,0		235	014	<b>U 1</b>	110	Deu	240
117																
118	Arg	Gln	Gln	Leu	Ile	Glu	Thr	Val	Glu	Tyr	Ser	Leu	Pro	Asn	Lys	His
119	_				245					250					255	
120																
121	Tyr	Phe	Glu		Asp	Leu	Ser	Trp		Lys	Gly	Leu	Gln	Asn	Thr	Gly
122				260					265					270		
123		_				_,		_		_	_	_	_		_	
124	Lys	Asn		Glu	Val	Phe	Ala		Gln	Ser	Asp	Pro		Gly	Leu	Ile
125 126			275					280					285			
127	Tva	Cva	Thr	Wa 1	Gly	Ara	Sor	Sor	T ou	Tva	Sor	Tve	Tau			
128	Dys	290	1111	vai	GIY	ALG	295	Ser	Leu	пåв	SEL	300	Dea			
129		2,0					2,5					500				
130	(2) INFO	RMAT:	ION I	FOR S	SEQ 3	D NO	0:2:									
131	• •				_											
132	(i)	SEQ	JENCI	E CHA	ARACI	reris	STICS	S :								
133		(A)	) LEI	NGTH:	: 302	2 ami	ino a	acida	3							
134		(B)	TYI	PE: a	amino	n aci	a									
		•														
135		(C	) STI	RANDI	EDNES	SS: 8	sing]	le								
136		(C	) STI	RANDI		SS: 8	sing]	le								
136 137	/ 4 4 \	(C)	) STI ) TOI	RANDI	EDNES GY: ]	SS: s Linea	sing] ar	le								
136 137 138	(ii)	(C)	) STI ) TOI	RANDI	EDNES GY: ]	SS: s Linea	sing] ar	le								
136 137 138 139		(C)	) STI ) TOI ECULI	RANDI POLOG E TYI	EDNES GY: ] PE: I	SS: s linea prote	sing] ar	le								
136 137 138 139 140	(ii) (iii)	(C)	) STI ) TOI ECULI	RANDI POLOG E TYI	EDNES GY: ] PE: I	SS: s linea prote	sing] ar	le								
136 137 138 139		(C) (D) MOLI	STI TOI	RANDI POLOG E TYI FICAI	EDNES GY: ] PE: p	SS: s linea prote	sing] ar	le								
136 137 138 139 140 141	(iii)	(C) (D) MOLI HYPO ORIO	STI TOI ECULI OTHE SINAI	RANDI POLOG E TYI FICAI L SOU	EDNES GY: ] PE: p	SS: s linea prote	sing] ar ein		flavi	18						
136 137 138 139 140 141 142	(iii)	(C) (D) MOLI HYPO ORIO	STI TOI ECULI OTHE SINAI	RANDI POLOG E TYI FICAI L SOU	EDNES GY: ] PE: I L: NO	SS: s linea prote	sing] ar ein		flavi	18						
136 137 138 139 140 141 142 143 144	(iii)	(C) (D) MOLI HYPO ORIC (A)	) STI ) TOI ECULI OTHET GINAL ) ORG	RANDI POLOG E TYI FICAI L SOU GANIS	EDNES GY: 1 PE: 1 L: NO URCE: SM: 1	SS: s linea prote C Nsper	sing] ar ein		flavı	ıs						
136 137 138 139 140 141 142 143 144 145	(iii) (vi)	(C) (D) MOLH HYPO ORIO (A) IMME	) STI ) TOI ECULE OTHE? GINAI ) ORG	RANDI POLOC E TYI FICAI L SOU GANIS	EDNES GY: 1 PE: 1 L: NO URCE: SM: 1	SS: s linea prote C Asper	sing] ar ein	lus 1		18						
136 137 138 139 140 141 142 143 144 145 146	(iii) (vi)	(C) (D) MOLH HYPO ORIO (A) IMME	) STI ) TOI ECULE OTHE? GINAI ) ORG	RANDI POLOC E TYI FICAI L SOU GANIS	EDNES GY: 1 PE: I L: NO URCE: SM: I	SS: s linea prote C Asper	sing] ar ein	lus 1		18						
136 137 138 139 140 141 142 143 144 145 146 147	(iii) (vi) (vii)	MOLI HYPO ORIC (A) IMME (B)	) STI ) TOI ECULH DTHET GINAI ) ORG EDIAT	RANDI POLOG TYI TICAI L SOU GANIS TE SO DNE:	EDNES GY: 1 PE: I L: NO URCE: Met-	SS: sinea prote Naper	singlar ein	lus 1	3e							
136 137 138 139 140 141 142 143 144 145 146 147 148	(iii) (vi)	MOLI HYPO ORIC (A) IMME (B)	) STI ) TOI ECULH DTHET GINAI ) ORG EDIAT	RANDI POLOG TYI TICAI L SOU GANIS TE SO DNE:	EDNES GY: 1 PE: I L: NO URCE: Met-	SS: sinea prote Naper	singlar ein	lus 1	3e							
136 137 138 139 140 141 142 143 144 145 146 147 148 149 150	(iii) (vi) (vii)	MOLI HYPO ORIO (A) IMME (B)	STI TOI ECULI OTHET GINAI ORG EDIAT CLO	RANDI POLOG E TYI FICAI L SOU GANIS TE SO DNE:	EDNES GY: ] PE: I L: NO URCE: Met-	SS: sinea prote PTION	singlar ein egill	lus i cidas	se O NO:	:2:	Lve	Aan	Acn	Val	Ara	Val
136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151	(iii) (vi) (vii) (xi) Met	MOLI HYPO ORIO (A) IMME (B)	STI TOI ECULI OTHET GINAI ORG EDIAT CLO	RANDI POLOG E TYI FICAI L SOU GANIS TE SO DNE:	EDNES GY: 1 PE: I  L: NO URCE: SM: I  OURCE Met- SCRII	SS: sinea prote Aspen	singlar ein egill	lus i cidas	se O NO:	:2: Gly	Lys	Asp	Asn	Val		Val
136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151	(iii) (vi) (vii)	MOLI HYPO ORIO (A) IMME (B)	STI TOI ECULI OTHET GINAI ORG EDIAT CLO	RANDI POLOG E TYI FICAI L SOU GANIS TE SO DNE:	EDNES GY: ] PE: I L: NO URCE: Met-	SS: sinea prote Aspen	singlar ein egill	lus i cidas	se O NO:	:2:	Lys	Asp	Asn	Val	Arg 15	Val
136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151	(iii) (vi) (vii) (xi) Met	MOLH HYPO ORIO (A) IMME (B) SEQU	STI TOI ECULE OTHET SINAI ORC EDIAT CLC JENCE	RANDI POLOG E TYI FICAI L SOU GANIS TE SO DNE:	EDNES GY: 1 PE: I  L: NO URCE: SM: A  OURCE Met- SCRIE Lys 5	SS: sinea prote PTION	singlar ein ein cgill ce ox	lus i kidas EQ II Arg	se ) NO: Tyr	:2: Gly 10					15	
136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154	(iii) (vi) (vii) (xi) Met	MOLH HYPO ORIO (A) IMME (B) SEQU	STI TOI ECULE OTHET SINAI ORC EDIAT CLC JENCE	RANDI POLOG E TYI FICAI L SOU GANIS TE SO DNE:	EDNES GY: 1 PE: I  L: NO URCE: SM: A  OURCE Met- SCRIE Lys 5	SS: sinea prote PTION	singlar ein ein cgill ce ox	lus i kidas EQ II Arg	se ) NO: Tyr	:2: Gly 10				Val Val 30	15	
136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156	(iii) (vi) (vii) (xi) Met 1 Tyr	MOLH HYPO ORIO (A) IMMH (B) SEQU Ser	) STI ) TOI ECULE OTHET GINAL ) ORG EDIAT ) CLG VAL	RANDI POLOG E TYI TICAI L SOU GANIS TE SC DNE: Val	EDNES GY: 1 PE: I L: NO URCE: SM: I OURCE Met- SCRII Lys 5	SS: sinea prote prote Napen -Urat PTION Ala Asp	singlar ein cgill ce ox N: SE Ala Glu	lus i cidas EQ II Arg Lys	Tyr Thr 25	:2: Gly 10 Gly	Val	Gln	Thr	Val 30	15 Tyr	Glu
136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156	(iii) (vi) (vii) (xi) Met 1 Tyr	MOLH HYPO ORIO (A) IMMH (B) SEQU Ser	STIP TO STIP T	RANDI POLOG E TYI TICAI L SOU GANIS TE SC DNE: Val	EDNES GY: 1 PE: I L: NO URCE: SM: I OURCE Met- SCRII Lys 5	SS: sinea prote prote Napen -Urat PTION Ala Asp	singlar ein cgill ce ox N: SE Ala Glu	lus i cidas EQ II Arg Lys	Tyr Thr 25	:2: Gly 10 Gly	Val	Gln	Thr Thr	Val	15 Tyr	Glu
136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156	(iii) (vi) (vii) (xi) Met 1 Tyr	MOLH HYPO ORIO (A) IMMH (B) SEQU Ser	) STI ) TOI ECULE OTHET GINAL ) ORG EDIAT ) CLG VAL	RANDI POLOG E TYI TICAI L SOU GANIS TE SC DNE: Val	EDNES GY: 1 PE: I L: NO URCE: SM: I OURCE Met- SCRII Lys 5	SS: sinea prote prote Napen -Urat PTION Ala Asp	singlar ein cgill ce ox N: SE Ala Glu	lus i cidas EQ II Arg Lys	Tyr Thr 25	:2: Gly 10 Gly	Val	Gln	Thr	Val 30	15 Tyr	Glu

#### Raw Sequence Listing

160 161	Lys	Ala 50	Asp	Asn	Ser	Val	Ile 55	Val	Ala	Thr	Asp	Ser 60	Ile	Lys	Asn	Thr
162		30					33					00				
163	Ile	Tyr	Ile	Thr	Ala	Lvs	Gln	Asn	Pro	Val	Thr	Pro	Pro	Glu	Leu	Phe
164	65	-1-				70					75					80
165																
166	Glv	Ser	Ile	Leu	Glv	Thr	His	Phe	Ile	Glu	Lvs	Tvr	Asn	His	Ile	His
167	•				85					90	-4	- 4 -			95	
168																
169	Ala	Ala	His	Val	Asn	Ile	Val	Сув	His	Arg	Trp	Thr	Arg	Met	Asp	Ile
170				100				-	105	_	_		_	110	_	
171																
172	Asp	Gly	Lys	Pro	His	Pro	His	Ser	Phe	Ile	Arg	Asp	Ser	Glu	Glu	Lys
173			115					120					125			
174																
175	Arg	Asn	Val	Gln	Val	Asp	Val	Val	Glu	Gly	Lys	Gly	Ile	Asp	Ile	Lys
176		130					135					140				
177																
178		Ser	Leu	Ser	Gly		Thr	Val	Leu	Lys		Thr	Asn	Ser	Gln	
179	145					150					155					160
180	_			_	_	_		_			_	_			_	
181	Trp	Gly	Phe	Leu		Asp	Glu	Tyr	Thr		Leu	Lys	Glu	Thr	_	Asp
182					165					170					175	
183	<b>3</b>	<b>-1</b> -	<b>T</b>		m1		**- 1	•		<b></b>		<b>01</b>	<b>.</b>	•	•	<b>5</b> 1
184	Arg	Ile	Leu		Thr	Asp	vaı	Asp		Thr	Trp	Gin	Trp		Asn	Phe
185 186				180					185					190		
187	S0=	C1	T 0	<b>61</b> m	<b>~1</b>	170 1	7	C	774 -	17-1	Desa	T	Dh.	3	21-	mh
188	Ser	Gly	195	GIII	GIU	Val	Arg	200	uis	vai	PIO	rya		Asp	ATG	Thr
189			193					200					205			
190	Trn	Ala	Thr	Ala	Ara	Glu	Va 1	Thr	T.011	T.sza	Thr	Dho	212	Glu	Aan	Agn
191		210			9	014	215		LCu	ביים	1	220		Olu	p	
192																
193	Ser	Ala	Ser	Val	Gln	Ala	Thr	Met	Tvr	Lvs	Met	Ala	Glu	Gln	Ile	Leu
194	225					230			-4-		235					240
195																
196	Ala	Arg	Gln	Gln	Leu	Ile	Glu	Thr	Val	Glu	Tyr	Ser	Leu	Pro	Asn	Lys
197		_			245					250	_				255	-
198																
199	His	Tyr	Phe	Glu	Ile	Asp	Leu	Ser	Trp	His	Lys	Gly	Leu	Gln	Asn	Thr
200				260					265					270		
201																
202	Gly	Lys		Ala	Glu	Val	Phe		Pro	Gln	Ser	Asp		Asn	Gly	Leu
203			275					280					285			
204				_	_	_										
205	Ile	Lys	Сув	Thr	Val	Gly		Ser	Ser	Leu	Lys		Lys	Leu		
206		290					295					300				
207					.=.											
208	(2) INFO	KMAT]	ON I	OR S	EQ 1	LD NC	):3:									
209	, , ,	ara.		. ~	. D. z. ~-		.m									
210 211	(1)	SEQU														
211				NGTH: PE: r				ILE								
414		(5)	TIF	e: I	iuCIE	TC 8	icta									

#### Raw Sequence Listing

04/15/92 09:05:43

215 216 (ii) MOLECULE TYPE: DNA (genomic) 217 218 219 (vii) IMMEDIATE SOURCE: 220 (B) CLONE: Preferred sequence for expression in 221 prokaryotes 222 223 224 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3: 225 226 ATGTCTGCGG TAAAAGCAGC GCGCTACGGC AAGGACAATG TTCGCGTCTA CAAGGTTCAC 60 227 228 AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG 120 229 230 GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC 180 231 232 ATTAAGAACA CCATTTACAT CACCGACAGC CAGACACCG TTACTCCTCC CGAGCTGTTC 240 233 234 GGCTCCATCC TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC 300 235 AACATTGTCT GCCACCGGTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC 360 236 AACATTGTCT GCCACCGGTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC 360 237 TTCATCCGCG ACAGCCAGAG GAAGCGGAA GTGCAGGTG ACGTGGTCGA GGGCAAGGGC 420 238 TTCATCCGCG ACAGCCAGAG GAAGCCGAA CTCCAGGAG GCAAGCCACA CCCTCACTCC 360 240 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGGTC 480 241 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGTCG 540 242 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGTCG 600 243 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG 600 244 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG 600 245 CACGTGCCTA AGTTCGATGC TACCTGGGC ACTGCTCGCG AGGTCACTCT GAAGACTTTT 660 246 GCGGCCAGC AGCTGATCGA GACTGTCGAG TACTCTGAC CTAACAAGCA CTATTTCGAA 780 247 ACCGACGTCG AGCTGATCGA GACTGTCGAG TACTCTGCC CTAACAAGCA CTATTTCGAA 780 248 CCCCAGTGCCAA GGCGCCCAAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 840 251 ACCGACCTGA GCTGGCCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 720 265 ACATTGC 267 CCTCAGTCGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 720 268 CCCCAGCACG AGCTGATCGA GACTGTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 720 267 CCTCAGTCGA GCTGGCCCAAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 720 268 CCCCAGCCAACGG TCTGATCAAAG TGTACCGTCG GCCGGTCCT TCTGAAGTCT 720 269 CCCCAGCCAACGG TCTGATCAAAG TGTACCGTCG GCCGGTCCT TCTGAAGTCT 720 260 (i) SEQUENCE CHARACTERISTICS: 261 (i) LENGTH: 906 base pairs	213	(C) STRANDEDNESS: single	
216 (ii) MOLECULE TYPE: DNA (genomic) 217 218 219 (vii) IMMEDIATE SOURCE: 220 (B) CLONE: Preferred sequence for expression in 221 prokaryotes 222 223 224 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3: 225 226 ATGTCTGCGG TAAAAGCAGC GCGCTACGGC AAGGACAATG TTCGCGTCTA CAAGGTTCAC 227 228 AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG 229 230 GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC 230 GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC 231 ATTAAGAACA CCATTTACAT CACCGCCAAG CAGAACCCG TTACTCCTCC CGAGCTGTTC 232 ATTAAGACA CCATTTACAT CACCGCCAAG CAGAACCCG TTACTCCTCC CGAGCTGTC 233 GGCTCCATCC TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC 233 AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC 236 AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC 237 238 TCCATCCCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC 240 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTCCTCA AGAGCACCAA CTCGCAGTTC 241 CGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC 242 ACCGACGTCG ATGCCACTTG GCAGTGGAGA AATTTCAGTG GACTCCAGGA GGTCCCTCG 243 ACCGACGTCG ATGCCACTTG GCAGTGGAGA AATTTCAGTG GACTCCAGGA GGTCCCTCG 244 ACCGACGTCG ATGCCACTTG GCAGTGGAGA AATTTCAGTG GACTCCAGGA GGTCCCTCG 245 ACCGTGCCTA AGTTCGATCC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT 246 CCTCAAGGATA ACAGTGCCAG GCTGCAGGAC ACTATGTACA AGATGGCAGA GCAAATCCTG 247 ACCGACGCCAA AGTTCCATGA GACTGTCGAG TACCTCTGAG GCCGTCCTC TACCTAGAC 248 ACCGACGCAA GGTGATCGA GACTGTCGAG TACCTCTGC CTAACAAGCA CTATTTCGAA 249 CCGCGCCAGC AGCTGATCGA GACTGTCGAG TACCTCTGC CTAACAAGCA CTATTTCGAA 251 ATCGACCTGA GCTGGCCAAA GGGCCTCCAA AACACCGGCA AGAACCCCAC CTATTTCGAA 252 ATCGACCTGA GCTGGCCAAA GGGCCTCCAA AACACCGGCA AGAACCCCAC GGTCTCTC TCTGAAGACC 253 ATCTGCACCTGA GCTGCTCCAA AACACCGGCA AGAACCCCAC CTATTTCGAA 253 ATCGACCTGA GCTGGCCCAACGG TCTGATCAAA AACACCGGCA AGAACCCCAC GGTCTCTC TCTGAAGTCT 254 CCTCAGTCGG ACCCCAACGG TCTGATCAAAG TGTACCGGCA AGAACCCCAC GGTCTCTC TCTGAAGTCT 255 AATTG 256 (2) INFORMATION FOR SEQ ID NO:4:	214	(D) TOPOLOGY: linear	
218 219 (vii) IMMEDIATE SOURCE: 220 (B) CLONE: Preferred sequence for expression in 221 prokaryotes 222 223 224 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3: 225 226 ATGTCTGCGG TAMAMGCAGC GCGCTACGGC AMGGACAATG TTCGCGTCTA CAMGGTTCAC 227 228 AMGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG 229 230 GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC 231 232 ATTAAGAACA CCATTTACAT CACCGACAGG CAGAACAGCG TCATTGTCGC CAGCTGTTC 240 231 232 AGCTCCATCC TGGGCACACA CTTCATTGAG AMGTACAACC ACATCCATGC CGCTCACGTC 235 236 AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGAGG GCAAGCCACA CCCTCACGTC 237 238 TTCATCCGCG ACAGCGAGGA GAAGCGGAT GTGCACGTGG ACGTGGTCGA GGGCAAGGGC 240 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTCTGA AGAGCACAA CTCGCAGTTC 241 TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC 242 ACCGACGTCG ATGCCACTG GCAGTGGAAGA AATTTCAGTG GACTCCAGGA GGTCCGCTCG 243 ACCGACGTCG ATGCCACTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCCTCG 244 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCCTCG 245 CACGTGCCTA AGTTCGATGC TACCTGGGC ACTGCTCGC AGGTCACTCT GAAGACTTTT 248 GCTGAAGATA ACAGTGCCAG GTACACCACA CTTAAGGAGA CCTGGGACGA GGTCCCTCG 249 240 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCCTCG 240 ACCGACGTCG AGCTCACTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCCTCG 241 CGCGCCCAC AGCTGATCGA GACTGTCGAG TACCTCGCC AGGTCACTCT GAAGACTTTT 248 GCTGAAGATA ACAGTGCCAG GCTGCAGGCC ACTATTTACA AGATGGCAGA GCAAATCCTG 249 240 CCCCAGCGCAG AGCTGATCGA GACTGTCCAG TACCTGCC CTAACAAGCA CTATTTCGAA 240 ACCGACGTCG AGCTGATCGA GACTGCTCGAG TACCTCGCC AGGTCACTCT GAAGACTTTT 248 GCTGAAGATA ACAGTGCCAG GCTGCAGGCC ACTATTTACA AGATGGCAGA GCAAATCCTG 249 240 ACCGACGCAGC AGCTGATCGA GACTGCTCGAG TACCTGCC CTAACAAGCA CTATTTCGAA 249 CCCCAGCAGC AGCTGATCGA GACTGCTCGC CTAACAAGCA CTATTTCGAA 240 ACCGACGCAACGG TTCGATCAAG TACCCGGCA AGAACCCCGA GGTCTCCT TCTGAAGTCT 247 248 CCTCAGTCGG ACCCAAACGG TTCGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 250 AATTG 251 ACCGACCTGA GCTGCACAA GGGCCTCCAA AACACCGGCA AGAACCCCGA GGTCTCCT TCTGAAGTCT 252 ACCGACCTAACGG TCTGATCAAG TGTACC			
218 219 (vii) IMMEDIATE SOURCE: 220 (B) CLONE: Preferred sequence for expression in 221 prokaryotes 222 223 224 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3: 225 226 ATGETCEGGG TAMAAGCAGG GCGCTACGGC AAGGACAATG TTCGCGTCTA CAAGGTTCAC 227 228 AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTG GCTTCTGGAG 229 230 GGTGAGATTG AGACCTCTTA CACCAAAGGCC GACAACAGGG TCATTGTCGC AACCGACTCC 230 GGTGAGATG AGACCTCTTA CACCACAGGC GACAACAGCG TCATTGTCGC AACCGACTCC 231 ATTAAGAACA CCATTTACAT CACCGCCAAG CAGAACCCCG TTACTCCTCC CGAGCTGTC 232 ATTAAGAACA CCATTTACAT CACCGCCAAG CAGAACCACG TCACTCCC CGAGCTGTC 233 GGCTCCATCC TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC 236 AACATTGTCT GCCACCGGTG GACCCGGATG GACATGACG GCAAGCCACA CCCTCACTCC 237 238 TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC 240 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGTGA AGAGCACCAA CTCGCAGTTC 241 TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC 242 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTCAGTG GACTCCAGGA GGTCCGCTCG 243 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTCAGTG GACTCCAGGA GGTCCGCTCG 244 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTCAGTG GACTCCAGGA GGTCCGCTCG 245 CACCGTCCT AGTTCGATGC TACCTGGGCC ACTGCTCGC AGGTCACTCT GAAGACTTTT 246 CCTGAAGATA ACAGTGCCAG CGTGCAGGC ACTATGTACA AGATGCCAGA GCAAATCCTG 247 248 CCTCAGTGCA AGCTGATCGA GACTGTCGAG TACCTGCG AGGTCACTCT GAAGACTTTT 247 248 CCTGAAGATA ACAGTGCCAG GGTCCACAA AACACCGGCA AGAACGCCGA GGTCTTCCCT 249 250 GCGCCCAGC AGCTGATCGA GACTGTCGAG TACCTGTGC CTAACAAGCA CTATTTCGAA 251 ATCCACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCCCT 252 ACCCACTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCCCT 253 ATCCACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCCCT 254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 255 AATTG 257 (2) INFORMATION FOR SEQ ID NO:4: 258 (1) SEQUENCE CHARACTERISTICS: 261 (1) SEQUENCE CHARACTERISTICS: 262 (1) SEQUENCE CHARACTERISTICS: 263 (1) TYPE: nucleic acid		(11) MOLECULE TYPE: DNA (genomic)	
(vii) IMMEDIATE SOURCE: (B) CLONE: Preferred sequence for expression in (CE) CLONE: Preferred sequence for expression in (CE) Prokaryotes  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3: (xi) SEQUENCE TAGAGCAGC GACGACAGC TTCCGGCTCTA CAAGGTTCAC (xi) SEQUENCE TAGAGCAGC TAGGACAGTA TTCCGGCTCTA CAAGGTTCACC (xi) SEQUENCE TAGAGCAGCAGT TACGAGCAGTA CAACCAGCTCACCC (xi) SEQUENCE TAGAGCAGCAGT TACGAGCAGTA CAACCAGCTCACCCC (xi) SEQUENCE TAGAGCAGCAGT TACGAGCAGTA CAACCAGCTCA CAACCAGCTC (xi) SEQUENCE TAGAGCAGCAGT TACGAGCAGTA CAACCAACCACCCCCCCCCC			
(B) CLONE: Preferred sequence for expression in prokaryotes  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:  ATGTCTGCGG TAAAAGCAGC GCGCTACGGC AAGGACAATG TTCGCGTCTA CAAGGTTCAC  AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG  GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC  ATTAAGAACA CCATTTACAT CACCGCCAAG CAGAACCACG TTACTCCTC CGAGCTGTTC  AACATTGTCT GCGCACCACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC  AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC  AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC  ACCACTCACTC TGCGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC  ACCACTCACTC TGCGCACACA GAACCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC  ACCACTCCCCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGCTCGA GGGCAAGGGC  ACCGACTCA AGTCGTCTCT GTCCGGCCTG ACCGTGTGA AGAGCACCAA CTCGCAGTTC  ACCGACGTCC AGGCCACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC  ACCGACGTCC ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG  ACCGACGTCC ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG  ACCGACGTCC ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG  CACCTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGC AGGTCACTCT GAAGACTTTT  CACCACGTCCTA AGTTCGATGC TACCTGGGCC ACTGCTGCG AGGTCACTCT GAAGACTTTT  CACCACCTGCA AGCTCGACAA GGCCTCCAAA AACACCGGCA AGAACGCCGA GCTCTTCCCT  ACCCACGTCCTA AGCTCGACAAA GGGCCTCCAAA AACACCGGCA AGAACGCCGA GCTCTTCCCT  ACCCACGTCCTA AGCTCGACAAA GGGCCTCCAAA AACACCGGCA AGAACGCCGA GCTCTTCCCT  ACCCACGTCCTA AGCTCGACAAA GGGCCTCCAAA AACACCGGCA AGAACGCCGA GCTCTTCCCT  ACCCACGTCCGA ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  CCTCAAGTCGA ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  AAATTG  CCTCAAGTCGA ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  AAATTG  CCTCAAGTCGA ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  AAATTG  CCTCAGTCCGC ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  AAATTG  ATTCATCCACCTCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  AAATTG  ATTCATCTCACACTCAAACCACACTACACACACACACA		4 111 7007777777	
221 prokaryotes  222 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:  223 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:  225 ATGTCTGCGG TAAAAGCAGC GCGCTACGGC AAGGACAATG TTCGCGTCTA CAAGGTTCAC  226 AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG  227 AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG  228 AATTAAGAACA CCATTTACAT CACCAAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC  230 GGTGAGATTG AGACCTCTTA CACCACAGGC CAGAACCCCG TTACTCCTCC CGAGCTGTTC  240 ATTAAGAACA CCATTTACAT CACCGCCAAG CAGAACCCCG TTACTCCTCC CGAGCTGTC  231 GGCTCCATCC TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC  236 AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC  237 TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC  239 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC  240 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC  241 TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAAGGAGA CCTGGGACCG TATCCTGAGC  242 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG  243 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG  244 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GCTCCAGTCT  245 CACCTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT  246 CACGTGCCTA AGTTCGATGC TACCTGGGC ACTGCTCGCG AGGTCACTCT GAAGACTTTT  247 GCTGAAGATA ACAGTGCCAG GCTGCAGGCC ACTGCTTCGC AGAACCCCGA GCTCTTCCCT  248 CCTCAAGCACA GCTGATCGA GACTGCTCGAG TACCTCTGC CTAACAAGCA CTATTTCGAA  250 ACCGCCCAG AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA  251 ACCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACCCCGA GGTCTTCCCT  252 ACCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACCCCGA GGTCTTCCCT  253 ACCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACCCCGA GGTCTTCCCT  254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGTCCTC TCTGAAGTCT  255 AAATTG  256 AAATTG  257 ATGTCTTCTCT  257 AATTTCAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC		, ,	
222 223 224 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:  225 226 ATGTCTGCGG TAAAAGCAGC GCGCTACGGC AAGGACAATG TTCGCGTCTA CAAGGTTCAC 60 227 228 AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG 120 229 230 GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC 180 231 ATTAAGAACA CCATTTACAT CACCGACAGG CAGAACACCC TTACTCCTCC CGAGCTGTC 240 232 AGGACTCC TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC 300 233 AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACGTC 360 234 ACCATTGACA AGTCGTCTCT GTCCGGCCTG ACCGTGTGA AGGACCCCA CCCTCACTCC 360 237 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGTGA AGAGCACCAA CTCGCAGTC 480 240 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGTGA AGAGCACCAA CTCGCAGTTC 480 241 ACCGACGTCG ATGCCACTTG GCAGTGGAAGA AATTCAGTG GACTCCAGGA GGTCCGCTCG 600 242 CACGACGTCG ATGCCACTTG GCAGTGGAAGA AATTCAGTG GACTCCAGGA GGTCCGCTCG 600 243 ACCGACGTCG ATGCCACTTG GCAGTGGAAGA AATTCAGTG GACTCCAGGA GGTCCGCTCG 600 244 CACGACGTCG ATGCCACTTG GCAGTGGAAGA AATTCAGTG GACTCCAGGA GGTCCCTCG 600 245 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT 660 246 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG 720 247 GCCGCCCAGC AGCTGATCGA GACTGCTCGAG TACCTCGTGC TAACAAGCA CTATTTCGAA 780 250 GCGCGCCAGC AGCTGATCGA GACTGCTCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 840 251 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 840 252 ACCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 840 253 AAATTG 906 254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 900 255 AAATTG 906 256 AAATTG 906 257 (2) INFORMATION FOR SEQ ID NO:4:		· · · · · · · · · · · · · · · · · · ·	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:  (xi) SEQUENCE CACAGCGC ACAGCGACAC TOCACGCC  (xi) SEQUENCE CACAGCGC ACAGCACACA TOCACACACACACACACACACACACACACACACACACACA		prokaryotes	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:  225 226 ATGTCTGCGG TARARGCAGC GCGCTACGGC AAGGACAATG TTCGCGTCTA CAAGGTTCAC 227 228 AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG 229 230 GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC 231 232 ATTARAGAACA CCATTTACAT CACCGACAG CAGAACCCCG TTACTCCTC CGAGCTGTTC 240 233 GGCTCCATCC TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC 236 AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC 237 238 TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC 240 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGTCA AGAGCACAA CTCGCAGTC 241 TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC 242 TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC 243 ACCGACGTCG ATGCCACTTG GCAGTGGAAGA AATTTCAGTG GACTCCAGGA GGTCCGCTCG 244 ACCGACGTCG ATGCCACTTG GCAGTGGAAGA AATTTCAGTG GACTCCAGGA GGTCCGCTCG 245 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCA AGAGCCCACA CTATTTCGAA 246 CACCGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCA AGATCCTC GAAGACTTTT 247 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATTTACA AGATGGCAGA GCAAATCCTG 248 GCCGCCCAGC AGCTGATCGA GACTGTCGAG TACCTGTGC TAACAAGCA CTATTTCGAA 249 GCCGCCCAGC AGCTGATCGA GACTGTCCAGA TACTCGTTGC CTAACAAGCA CTATTTCGAA 240 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 241 ACCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 242 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 244 ACCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 245 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 246 AAATTG 257 (2) INFORMATION FOR SEQ ID NO:4: 258 (1) SEQUENCE CHARACTERISTICS: 259 (1) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid			
ATTANGANCA CCATTTACAT CACCGCAAG GAAGCACAC ACTACGCA ACCGTCCTCA CAAGGTTCAC  ACCGACCGC ACCGCCAACGC GACACCACC ACCGCCCAA ACCCCCCCAACCACCC CACCGCCCCCCCC		(vi) SPOUPNCE DESCRIPTION. SPO ID NO. 2.	
ATGTCTGCGG TANANGCAGC GCGCTACGGC AAGGACAATG TTCGCGTCTA CAAGGTTCAC  AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG  AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG  GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC  180  ATTAAGAACA CCATTTACAT CACCGCCAAG CAGAACCCCG TTACTCCTCC CGAGCTGTTC  AACATTGTCT GGCCACCGCTG GACCCGGATG AAGTACAACC ACATCCATGC CGCTCACGTC  AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC  AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC  ACCGTCATCC TGCGAGAGA GAAGCGGAAT GTGCAGGTG ACGTGCTCGA GGGCAAGGGC  ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACAA CTCGCAGGTC  ACCGACGTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC  ACCGACGTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC  ACCGACGTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC  CACCGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGC AGGTCACTCT GAAGACTTTT  ACCGACGTCC AGGTCGAGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT  ACCGACGTCC AGGTCGACGA GACCGTCGAGC ACTATGTACA AGATGGCAGA GCAAATCCTG  GCTGAAGATA ACAGTGCCAG GCTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG  GCGCGCCCAG AGCTGATCGA GACTGTCGAG TACCCTTGC CTAACAAGCA CTATTTCGAA  ACCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGCA AGAACGCCGA GGTCTTCGCT  AACAGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGCA AGAACGCCGA GGTCTTCGCT  AACAGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGCA AGAACGCCGA GGTCTTCGCT  AACAGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT  AACTGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT  AACATTG  (2) INFORMATION FOR SEQ ID NO:4:  (3) SEQUENCE CHARACTERISTICS:  (4) LENGTH: 906 base pairs  (5) GLGCGCCACGAGG CAACCGCAACGG TCTGCTCGCG GCGGTCCTC TCTGAAGTCA  (B) TYPE: nucleic acid		(XI) SEQUENCE DESCRIPTION: SEQ ID NO:3:	
AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG  229 230 GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC  231 ATTAAGAACA CCATTTACAT CACCGCCAAG CAGAACCCCG TTACTCCTCC CGAGCTGTTC  240 232 AGCATTGTCT TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC  235 ACCATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC  236 AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC  237 TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC  239 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC  240 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC  241 TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC  242 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG  243 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG  244 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG  245 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT  246 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG  247 GCCGCCCAGC AGCTGATCGA GACTGTCCGAG TACCTGTGC CTAACAAGCA CTATTTCGAA  250 GCCGCCCAGC AGCTGATCGA GACTGTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT  251 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT  252 ACCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  253 (2) INFORMATION FOR SEQ ID NO:4:  254 (1) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 906 base pairs  (B) TYPE: nucleic acid		ATCTCTCCCC TAAAACCACC CCCCTACCCC AACCACAATC TTCCCCTCTA CAACCTTCAC	60
AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG  GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC  ATTAAGAACA CCATTTACAT CACCGCCAAG CAGAACCCCG TTACTCCTCC CGAGCTGTTC  GGCTCCATCC TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC  AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC  AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC  TCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC  TCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC  TCATCCGCG ACAGCGAGGA GAACCGCAT GTCCAGGTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC  TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC  TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC  TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC  ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG  CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT  GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT  GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG  GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA  ACCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT  AACATTG  CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  900  CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  900  CTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  (A) LENGTH: 906 base pairs  (B) TYPE: nucleic acid		MIGIGIOGO HERROCAGO GOOGIAGGO MAGAACMIG IIGGGGIGIA CAMGGIICAG	00
229 230 GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC 231 232 ATTAAGAACA CCATTTACAT CACCGCCAAG CAGAACCCCG TTACTCCTCC CGAGCTGTTC 240 233 GGCTCCATCC TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC 235 236 AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCCAC CCCTCACCTC 237 238 TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC 239 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC 240 ACCGACGTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC 241 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG 242 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCACGAG GGTCCGCTCG 243 ACCGACGTCA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT 244 ACCGACGTCA AGTTCGATGC TACCTGGGCC ACTACTCACGA GCAAATCCTG 245 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTACTTACA AGATGGCAGA GCAAATCCTG 246 GCGCACCAG AGCTGATCGA GACTGTCGAG TACCTGTTCC CTAACAAGCA CTATTTCGAA 247 ACCGACGTCA AGCTGATCGA GACTGTCGAG TACCTGTTCC CTAACAAGCA CTATTTCGAA 248 GCGCGCCAGC AGCTGATCGA GACTGTCCGAG TACCTGTTCC CTAACAAGCA CTATTTCGAA 249 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACCCGTTCC CTAACAAGCA CTATTTCGAA 250 ACCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCCCT 251 ATCGACCTGA GCCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 252 ATCGACCTGA GCCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 253 AAATTG 254 (2) INFORMATION FOR SEQ ID NO:4: 255 (2) (3) SEQUENCE CHARACTERISTICS: (4) LENGTH: 906 base pairs (5) TYPE: nucleic acid		AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG	120
230 GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC 231 232 ATTAAGAACA CCATTTACAT CACCGCCAAG CAGAACCCCG TTACTCCTCC CGAGCTGTTC 240 233 234 GGCTCCATCC TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC 300 235 236 AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC 360 237 238 TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC 239 240 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC 480 241 242 TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC 243 244 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG 245 246 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGC AGGTCACTCT GAAGACTTTT 247 248 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG 249 250 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACCTGTGC CTAACAAGCA CTATTTCGAA 251 ATCGACCTGA GCTGGCACAA GGCCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 252 253 ACCGACCTGA GCTGGCACAA GGCCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 254 255 AAATTG 257 258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (1) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 261 (B) TYPE: nucleic acid			
231 232 ATTANGARCA CCATTTACAT CACCGCCAAG CAGAACCCCG TTACTCCTCC CGAGCTGTTC 240 233 234 GGCTCCATCC TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC 300 235 236 AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC 360 237 238 TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTG ACGTGGTCGA GGGCAAGGGC 239 240 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC 480 241 242 TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC 243 244 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG 245 246 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT 248 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG 249 250 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA 251 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACCCCGA GGTCTTCGCT 252 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACCCCGA GGTCTTCGCT 253 254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 255 AAATTG 256 (1) SEQUENCE CHARACTERISTICS: 257 (2) INFORMATION FOR SEQ ID NO:4: 258 (1) SEQUENCE CHARACTERISTICS: 259 (1) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid		GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC	180
233 234 GGCTCCATCC TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC 300 235 236 AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC 360 237 238 TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC 420 239 240 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACAA CTCGCAGTTC 480 241 242 TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC 540 243 244 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG 600 245 246 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT 660 247 248 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG 249 250 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA 780 251 252 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 840 253 254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 900 255 AAATTG 906 261 SEQUENCE CHARACTERISTICS: (A) LENGTH: 906 base pairs (B) TYPE: nucleic acid			
GCTCCATCC TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC  300 325 326 AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC 360 327 238 TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC 329 240 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC 480 421 242 TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC 423 244 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG 450 451 452 453 GCTGAAGATA ACAGTGCCAC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT 460 475 476 GCTGAAGATA ACAGTGCCAG CGTGCAGGC ACTATGTACA AGATGGCAGA GCAAATCCTG 477 478 GCTGAAGATA ACAGTGCCAG GGTCGTCGAG TACCTGTGC CTAACAAGCA CTATTTCGAA 478 479 480 490 491 492 493 494 495 495 496 497 497 498 499 499 409 409 409 409 409 409 409 409	232	ATTAAGAACA CCATTTACAT CACCGCCAAG CAGAACCCCG TTACTCCTCC CGAGCTGTTC	240
235 236 AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC 237 238 TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC 239 240 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC 241 242 TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC 243 244 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG 245 246 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGC AGGTCACTCT GAAGACTTTT 248 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG 249 250 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACCTGGTGC CTAACAAGCA CTATTTCGAA 251 252 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 253 254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 255 256 AAATTG 257 258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid	233		
AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC  360  237  238 TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC  420  240 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC  480  241  242 TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC  243  244 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG  245  246 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT  460  247  248 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG  249  250 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA  780  251  ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT  840  253  254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  906  257  258 (2) INFORMATION FOR SEQ ID NO:4:  (i) SEQUENCE CHARACTERISTICS:  (ii) SEQUENCE CHARACTERISTICS:  (ii) SEQUENCE CHARACTERISTICS:  (iii) LENGTH: 906 base pairs  (iiii) TYPE: nucleic acid	234	GGCTCCATCC TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC	300
TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC 420 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC 480 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC 480 TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC 540 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG 600 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT 660 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT 720 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA 780 CTCAGTCGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 840 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 900 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 906 AAATTG 906 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 906 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACAGTCT GCTCAAACTCTGTCT 906 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACAGTCT GCTCAAACTCTGTCT 906 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACAGTCT GCTCAAACTCTTTTTCAACAGTCTTTTTCAACAGTCTTTTTTTT	235		
TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC  TCCCCCCCCCCCCCCCCCCCCCCCCCCC	236	AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC	360
239 240 ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC 241 242 TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC 243 244 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG 245 246 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT 247 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG 248 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG 249 250 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA 251 252 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 253 254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 255 256 AAATTG 257 258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid			
ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC  241  242 TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC  243  244 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG  245  246 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT  248 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG  249  250 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA  251  252 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT  253  254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  255  256 AAATTG  (1) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 906 base pairs  (B) TYPE: nucleic acid		TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC	420
TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC  TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC  ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG  CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT  CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG  CACGTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG  CACGTGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA  TRO  CACGTGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA  TRO  CACGTGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA  TRO  CACGTGCACACAGA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT  CACGTGCGACCACAGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  CACGTGCACACAGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  CACGTGCACACACAGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  CACGTGCACACACAGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  CACGTGCACACACAGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  CACGTGCACACACACAGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  CACGTGCACACACACAGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  CACGTGCACACACACAGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  CACGTGCACACACACACACACACACACACACACACACACA			
TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC  TGCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG  ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG  CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT  GCCGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG  GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA  T80  TCCCCCAGCCAGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT  ACCCCCAGCTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT  TCCCCCCCCCCCCCCCCCCCCCCCCCCC		ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC	480
243 244 ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG 600 245 246 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT 660 247 248 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG 720 249 250 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA 780 251 252 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 840 253 254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 900 255 256 AAATTG 906 257 258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid		MOCCOCOMMON MOCCOMO CON CONTROL CONTRO	- 40
ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG  245  246 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT 660  247  248 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG 720  250 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA 780  251 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 840  253 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 900  255 AAATTG 906  260 (i) SEQUENCE CHARACTERISTICS:  261 (A) LENGTH: 906 base pairs  262 (B) TYPE: nucleic acid		TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC	540
245 246 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT 660 247 248 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG 720 249 250 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA 780 251 252 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 840 253 254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 900 255 256 AAATTG 906 257 258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid		ACCCACCTOC ATTCCCACTTC COACTCCAAC AATTTTCACTC CACTCCAACCA CCTTCCCCTCC	600
246 CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT 660 247 248 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG 720 249 250 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA 780 251 252 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 840 253 254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 900 255 256 AAATTG 906 257 258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid		ACCONCING AIGCCACITY GCASTGGAAG AATTICAGIY GACTCCAGGA GGTCCGCTCG	800
247 248 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG 249 250 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA 251 252 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 253 254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 255 256 AAATTG 2906 257 258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid		CACCTCCCTA ACTTCCATCC TACCTCCCCC ACTCCCCC ACCTCACTCT CAACACTTTT	660
248 GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG 249 250 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA 251 252 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 253 254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 255 256 AAATTG 2906 257 258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid		ONDOTOCOTA NOTICONICO INCOLOGOCO NOTICOCO NOTICACION GARGACITI	000
249 250 GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA 780 251 252 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 840 253 254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 900 255 256 AAATTG 906 257 258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid		GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG	720
251 252 ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT 253 254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 255 256 AAATTG 257 258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid			
ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT  253 254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 255 256 AAATTG 257 258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid	250	GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA	780
253 254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 255 256 AAATTG 257 258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid	251		
254 CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT 255 256 AAATTG 257 258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid	252	ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT	840
255 256 AAATTG 257 258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid	253		
256 AAATTG 257 258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid		CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT	900
257 258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid			
258 (2) INFORMATION FOR SEQ ID NO:4: 259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid		AAATTG	906
259 260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid			
260 (i) SEQUENCE CHARACTERISTICS: 261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid		(2) INFORMATION FOR SEQ ID NO:4:	
261 (A) LENGTH: 906 base pairs 262 (B) TYPE: nucleic acid			
262 (B) TYPE: nucleic acid		• • -	
263 (C) STRANDERNESS, Girala	263	(B) TYPE: Nucleic acid (C) STRANDEDNESS: single	
264 (D) TOPOLOGY: linear			
265		(b) Totoboot: Timedi	

#### Raw Sequence Listing

04/15/92 09:05:49

266 267	(ii) MOLECULE TYPE: DNA (genomic)	
268		
269	(vii) IMMEDIATE SOURCE:	
270	(B) CLONE: Preferred sequence for expression in	
271	eukaryotes	
272		
273		
274	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:	
275	. , -	
276	ATGTCTGCTG TTAAGGCTGC TAGATACGGT AAGGACAACG TTAGAGTCTA CAAGGTTCAC	60
277		
278	AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG	120
279		
280	GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC	180
281		
282	ATTAAGAACA CCATTTACAT CACCGCCAAG CAGAACCCCG TTACTCCTCC CGAGCTGTTC	240
283		240
284	GGCTCCATCC TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC	300
285	decision independent circuitana andiacance acatecanoe enciconesic	300
286	AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC	360
287	ARCHITOTOL GCCACCGCIG GACCCGGALG GACALIGACG GCAAGCCACA CCCICACICC	300
288	TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC	420
289	TICATECECE ACAGEGRAGA GARGEGRAFI GIGERGGIGG ACGIGGIEGA GGGERAGGGE	420
290	ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC	480
291	ATCOMITATION AGTOGRACIO GEOGGEORGE ACCORDIGENCEM CECGOCAGETE	400
292	TOCOCCOTTOC TOCOCTOROS CTROSPOROS CTROSPOROS CONTRACTOROS	E40
293	TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC	540
294	ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG	600
295	ACCORCICG AIGCCACIIG GCAGIGGAAG AAIIICAGIG GACICCAGGA GGICCGCICG	600
296	CACCECCERA ACEMECANICA MACAMACACA ACMACACA ACEMAACMAM CAACACMAMM	660
297	CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT	660
298	COMONNON NONCOCORO COMOCNOCO NONNOCONON NONCOCONON CONNINCOCO	720
299	GCTGAAGATA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG	720
	000000000 2000000000 020000000 m10000000 0011011000 0011011000	500
300	GCGCGCCAGC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA	780
301	3.0003.00003.000003.013.00000003.013.0000003.00000000	
302	ATCGACCTGA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT	840
303		
304	CCTCAGTCGG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT	900
305		
306	AAATTG	906
307		
308	(2) INFORMATION FOR SEQ ID NO:5:	
309		
310	(i) SEQUENCE CHARACTERISTICS:	
311	(A) LENGTH: 14 base pairs	
312	(B) TYPE: nucleic acid	
313	(C) STRANDEDNESS: single	
314	(D) TOPOLOGY: linear	
315		
316	(ii) MOLECULE TYPE: DNA (genomic)	
317		
318	(iii) HYPOTHETICAL: NO	

#### Raw Sequence Listing

04/15/92 09:05:52

319 320		
321	(vii) IMMEDIATE SOURCE:	
322	(VII) INTEDIATE SOURCE:  (B) CLONE: Preferred non-translated 5' sequence for	
323	animal cells	
324	animal cells	
325		
326	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:	
327		
328	AGCTTGCCGC CACT	14
329		
330	(2) INFORMATION FOR SEQ ID NO:6:	
331		
332	(i) SEQUENCE CHARACTERISTICS:	
333	(A) LENGTH: 906 base pairs	
334	(B) TYPE: nucleic acid	
335	(C) STRANDEDNESS: double	
336	(D) TOPOLOGY: linear	
337		
338	(ii) MOLECULE TYPE: DNA (genomic)	
339 340	(111) HUDOMHERTONI - NO	
341	(iii) HYPOTHETICAL: NO	
342		
343	(vii) IMMEDIATE SOURCE:	
344	(B) CLONE: Preferred sequence for expression in animal	
345	cells	
346	***************************************	
347		
348	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:	
349		
350	ATGTCCGCAG TAAAAGCAGC CCGCTACGGC AAGGACAATG TCCGCGTCTA CAAGGTTCAC	60
351		
352	AAGGACGAGA AGACCGGTGT CCAGACGGTG TACGAGATGA CCGTCTGTGT GCTTCTGGAG	120
353		
354	GGTGAGATTG AGACCTCTTA CACCAAGGCC GACAACAGCG TCATTGTCGC AACCGACTCC	180
355		
356	ATTAAGAACA CCATTTACAT CACCGCCAAG CAGAACCCCG TTACTCCTCC CGAGCTGTTC	240
357	CCCTCC1500 BCCCC1510 CDCC1510 BCCC	
358 359	GGCTCCATCC TGGGCACACA CTTCATTGAG AAGTACAACC ACATCCATGC CGCTCACGTC	300
360	AACAMMCMCM CCCACCCCCC CACCCCAMC CACAMMCACC CCAACCCACA CCCMCACMCA	200
361	AACATTGTCT GCCACCGCTG GACCCGGATG GACATTGACG GCAAGCCACA CCCTCACTCC	360
362	TTCATCCGCG ACAGCGAGGA GAAGCGGAAT GTGCAGGTGG ACGTGGTCGA GGGCAAGGGC	420
363	TICHICCOCG ACAGCGAGGA GAAGCGGAAT GIGCAGGIGG ACGIGGICGA GGGCAAGGGC	420
364	ATCGATATCA AGTCGTCTCT GTCCGGCCTG ACCGTGCTGA AGAGCACCAA CTCGCAGTTC	480
365	middlinia. Moldololol bloodoolo Modolololi Admocaccaa Clogoagilo	400
366	TGGGGCTTCC TGCGTGACGA GTACACCACA CTTAAGGAGA CCTGGGACCG TATCCTGAGC	540
367	221111111111111111111111111111111111111	230
368	ACCGACGTCG ATGCCACTTG GCAGTGGAAG AATTTCAGTG GACTCCAGGA GGTCCGCTCG	600
369		
370	CACGTGCCTA AGTTCGATGC TACCTGGGCC ACTGCTCGCG AGGTCACTCT GAAGACTTTT	660
371		

#### Raw Sequence Listing

04/15/92 09:05:58

372	GCTGAAGA	TA ACAGTGCCAG CGTGCAGGCC ACTATGTACA AGATGGCAGA GCAAATCCTG	720
373			
374	GCGCGCCA	GC AGCTGATCGA GACTGTCGAG TACTCGTTGC CTAACAAGCA CTATTTCGAA	780
375			
376	ATCGACCT	GA GCTGGCACAA GGGCCTCCAA AACACCGGCA AGAACGCCGA GGTCTTCGCT	840
377			
378	CCTCAGTC	GG ACCCCAACGG TCTGATCAAG TGTACCGTCG GCCGGTCCTC TCTGAAGTCT	900
379			
380	AAATTG		906
381			
382	(2) INFO	RMATION FOR SEQ ID NO:7:	
383			
384	(i)	SEQUENCE CHARACTERISTICS:	
385		(A) LENGTH: 23 base pairs	
386		(B) TYPE: nucleic acid	
387		(C) STRANDEDNESS: single	
388		(D) TOPOLOGY: linear	
389			
390	(ii)	MOLECULE TYPE: DNA (genomic)	
391			
392	(iii)	HYPOTHETICAL: NO	
393			
394			
395	(A11)	IMMEDIATE SOURCE:	
396		(B) CLONE: reverse transcription primer	
397			
398	4	CHAURNAR RECORDERAN. ORG. TR. NO. 7.	
399 400	(XI)	SEQUENCE DESCRIPTION: SEQ ID NO:7:	
401	CATCCCCC	CC CTTTTTTTT TTT	23
402	GAICCGGG	· ·	23
403	(2) INFO	RMATION FOR SEQ ID NO:8:	
404	(2) 1010	MATION FOR SEQ ID NO.0.	
405	(3)	SEQUENCE CHARACTERISTICS:	
406	(-)	(A) LENGTH: 10 amino acids	
407		(B) TYPE: amino acid	
408		(C) STRANDEDNESS: single	
409		(D) TOPOLOGY: linear	
410		(-,	
411	(ii)	MOLECULE TYPE: peptide	
412	<b>\</b> ,		
413	(iii)	HYPOTHETICAL: NO	
414	, ,	·	
415			
416	(vii)	IMMEDIATE SOURCE:	
417	•	(B) CLONE: Hydrolysis product T 17	
418			
419			
420	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:8:	
421	•		
422	Asn	Val Gln Val Asp Val Val Glu Gly Lys	
423	1	5 10	
424			

425	(2) INFORMATION FOR SEQ ID NO:9:
426	
427	(i) SEQUENCE CHARACTERISTICS:
428	(A) LENGTH: 8 amino acids
429	(B) TYPE: amino acid
430	(C) STRANDEDNESS: single
431	(D) TOPOLOGY: linear
432	
433	(ii) MOLECULE TYPE: peptide
434	
435	(iii) HYPOTHETICAL: NO
436	
437	
438	(vii) IMMEDIATE SOURCE:
439	(B) CLONE: Hydrolysis product T 20
440	
441	
442	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:
443	
444	Asn Phe Ser Gly Leu Gln Glu Val
445	1 5
446	
447	(2) INFORMATION FOR SEQ ID NO:10:
448	-
449	(i) SEQUENCE CHARACTERISTICS:
450	(A) LENGTH: 6 amino acids
451	(B) TYPE: amino acid
452	(C) STRANDEDNESS: single
453	(D) TOPOLOGY: linear
454	` ,
455	(ii) MOLECULE TYPE: peptide
456	* *
457	(iii) HYPOTHETICAL: NO
458	,
459	
460	(vii) IMMEDIATE SOURCE:
461	(B) CLONE: Hydrolysis product T 23
462	(=, ===================================
463	
464	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:
465	(000) 0020000000000000000000000000000000
466	Phe Asp Ala Thr Trp Ala
467	1 5
468	-
469	(2) INFORMATION FOR SEQ ID NO:11:
470	(2) 1.11 01.11 101. 101. 102 10 110.111.
471	(i) SEQUENCE CHARACTERISTICS:
472	(A) LENGTH: 8 amino acids
473	(B) TYPE: amino acid
474	(C) STRANDEDNESS: single
475	(D) TOPOLOGY: linear
476	(b) totoboot. Itheat
477	(ii) MOLECULE TYPE: peptide
<b>4</b> //	(II) HOUROODS IIES: Pahrida

478		
479	(iii)	HYPOTHETICAL: NO
480		
481		
482	(vii)	IMMEDIATE SOURCE:
483	, ,	(B) CLONE: Hydrolysis product T 27
484		(=, ====== u, ==========================
485		
486	(vi)	SEQUENCE DESCRIPTION: SEQ ID NO:11:
487	(71)	DIQUINOS DISONITITON. DEQ 15 NO.11.
488	ніа	Tyr Phe Glu Ile Asp Leu Ser
489	1	5
490	•	•
491	/2) TNEO	RMATION FOR SEQ ID NO:12:
492	(2) INFO	AMAIION FOR SEQ ID NO:12:
493	/: \	CECHENGE CUADACMEDICATOS.
494	(1)	SEQUENCE CHARACTERISTICS:
494		(A) LENGTH: 13 amino acids
		(B) TYPE: amino acid
496		(C) STRANDEDNESS: single
497		(D) TOPOLOGY: linear
498		
499	(ii)	MOLECULE TYPE: peptide
500		
501	(iii)	HYPOTHETICAL: NO
502		
503		
504	(vii)	IMMEDIATE SOURCE:
505		(B) CLONE: Hydrolysis product T 28
506		· ·
507		
508	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:12:
509	(/	
510	Ile	Leu Ser Thr Asp Val Asp Ala Thr Trp Gln Trp Lys
511	1	5 10
512	-	10
513	(2) INFO	RMATION FOR SEQ ID NO:13:
514	(2) INFO	MARION FOR BEQ ID NO.13.
515	/33	SEQUENCE CHARACTERISTICS:
516	( - )	(A) LENGTH: 11 amino acids
517		·
		(B) TYPE: amino acid
518		(C) STRANDEDNESS: single
519		(D) TOPOLOGY: linear
520		
521	(ii)	MOLECULE TYPE: peptide
522		
523	(iii)	HYPOTHETICAL: NO
524		
525		
526	(vii)	IMMEDIATE SOURCE:
527		(B) CLONE: Hydrolysis product T 29
528		
529		
530	(xi)	SEQUENCE DESCRIPTION: SEO ID NO:13:

531		
532	His	Tyr Phe Glu Ile Asp Leu Ser Trp His Lys
533	1	5 10
534		
535	(2) INFO	RMATION FOR SEQ ID NO:14:
536	(_,	
537	(1)	SEQUENCE CHARACTERISTICS:
538	( - )	(A) LENGTH: 11 amino acids
539		(B) TYPE: amino acid
540		(C) STRANDEDNESS: single
541		· · ·
542		(D) TOPOLOGY: linear
543	(22)	NOT FOUL F. MVDF
	(11)	MOLECULE TYPE: peptide
544		***************************************
545	(111)	HYPOTHETICAL: NO
546		
547		
548	(vii)	IMMEDIATE SOURCE:
549		(B) CLONE: Hydrolysis product T 31
550		
551		
552	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:14:
553		
554	Ser	Thr Asn Ser Gln Phe Trp Gly Phe Leu Arg
555	1	5 10
556		
557	(2) INFO	RMATION FOR SEQ ID NO:15:
558		
559	(i)	SEQUENCE CHARACTERISTICS:
560	, ,	(A) LENGTH: 16 amino acids
561		(B) TYPE: amino acid
562		(C) STRANDEDNESS: single
563		(D) TOPOLOGY: linear
564		(-) 10102001 22012
565	(11)	MOLECULE TYPE: peptide
566	()	
567	/1111	HYPOTHETICAL: NO
568	(111)	miroinblicab. No
569		
570	/*** i \	IMMEDIATE SOURCE:
571	( ^ TT )	
572		(B) CLONE: Hydrolysis product T 32
573	4	40410145
574	(XI)	SEQUENCE DESCRIPTION: SEQ ID NO:15:
575		
576		Asn Pro Val Thr Pro Pro Glu Leu Phe Gly Ser Ile Leu Gly Th
577	1	5 10 15
578		
579		
580	(2) INFO	RMATION FOR SEQ ID NO:16:
581		
582	(i)	SEQUENCE CHARACTERISTICS:
583		(A) LENGTH: 16 amino acids

584		(B) TYPE: amino acid
585		(C) STRANDEDNESS: single
586		(D) TOPOLOGY: linear
587		
588	(11)	MOLECULE TYPE: peptide
589		
590	(iii)	HYPOTHETICAL: NO
591		
592		
593	(vii)	IMMEDIATE SOURCE:
594		(B) CLONE: Hydrolysis product T 33
595		
596		
597	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:16:
598		
599	Gln	Asn Pro Val Thr Pro Pro Glu Leu Phe Gly Ser Ile Leu Gly Thr
600	1	5 10 15
601		
602		
603	(2) INFO	RMATION FOR SEQ ID NO:17:
604		
605	(i)	SEQUENCE CHARACTERISTICS:
606		(A) LENGTH: 17 amino acids
607		(B) TYPE: amino acid
608		(C) STRANDEDNESS: single
609		(D) TOPOLOGY: linear
610		
611	(ii)	MOLECULE TYPE: peptide
612		• •
613	(iii)	HYPOTHETICAL: NO
614	, ,	
615		
616	(vii)	IMMEDIATE SOURCE:
617	• • •	(B) CLONE: Hydrolysis product V 1
618		(=, ===================================
619		
620	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:17:
621	(/	
622	Tvr	Ser Leu Pro Asn Lys His Tyr Phe Glu Ile Asp Leu Ser Trp His
623	1	5 10 15
624	-	
625	Lys	
626	-1-	
627		
628	(2) INFO	RMATION FOR SEQ ID NO:18:
629	(2) 11110	AMITON FOR ODG ID NO.10.
630	/ 1 1	SEQUENCE CHARACTERISTICS:
631	(-)	(A) LENGTH: 16 amino acids
632		(B) TYPE: amino acid
633		(C) STRANDEDNESS: single
634		(D) TOPOLOGY: linear
635		(b) 1010b001. Illieat
636	/ 1 1 1	MOLECULE TYPE: peptide
JJ0	( + + )	MODECODE IIFE: Pepcide

637		
638	(iii)	HYPOTHETICAL: NO
639		
640		
641	(vii)	IMMEDIATE SOURCE:
642		(B) CLONE: Hydrolysis product V 2
643		
644		
645	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:18:
646	••- •	
647		Thr Leu Lys Thr Phe Ala Glu Asp Asn Ser Ala Ser Val Gln Ala
648	1	5 10 15
649 650		
651	(2) INFO	DUATION FOR CRO ID NO.10.
652	(2) INFO	RMATION FOR SEQ ID NO:19:
653	/53	SEQUENCE CHARACTERISTICS:
654	(+)	(A) LENGTH: 24 amino acids
655		(B) TYPE: amino acid
656		(C) STRANDEDNESS: single
657		(D) TOPOLOGY: linear
658		(5) 101020011 12002
659	(ii)	MOLECULE TYPE: peptide
660	, ,	
661	(iii)	HYPOTHETICAL: NO
662		
663		
664	(vii)	IMMEDIATE SOURCE:
665		(B) CLONE: Hydrolysis product V 3
666		
667		
668	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:19:
669		
670		Ser Tyr Thr Lys Ala Asp Asn Ser Val Ile Val Asp Thr Asp Ser
671	1	5 10 15
672	<b>-</b> 1 -	Tue 1 - Mb - Tl - M - Tl - M
673 674	116	Lys Asn Thr Ile Tyr Ile Thr
675		20
676	/2\ TNEOI	DANTION FOR SEC ID NO.20.
677	(2) INFO	RMATION FOR SEQ ID NO:20:
678	(1)	SEQUENCE CHARACTERISTICS:
679	(+)	(A) LENGTH: 28 amino acids
680		(B) TYPE: amino acid
681		(C) STRANDEDNESS: single
682		(D) TOPOLOGY: linear
683		\-, <del></del>
684	(ii)	MOLECULE TYPE: peptide
685	, ,	• • • • • • • • • • • • • • • • • •
686	(iii)	HYPOTHETICAL: NO
687	•	
688		
689	(vii)	IMMEDIATE SOURCE:

690		(B)	CLONE	Hyd	roly	sis j	prod	uct '	V 5						
691															
692	/ <del>!</del>	CROU		300D T	DM T 0.1		- T								
693 694	(XI)	SEQUE	ENCE D	SCRI	PTIO	N: 5.	eð Ti	טא כ	:20:						
695	<b>C1</b>	T (		. 7	T1.	T	Co	Com	T	C	<b>01</b>	T	mh	17.01	T
		гур (	Gly Ile		116	гля	ser	ser		ser	GIY	Leu	THE		Leu
696	1			5					10					15	
697	T	C [	nh n		<b>01</b> -	nh-	m	<b>a</b> 1	Dh.	<b>T</b>	3				
698 699	гув	ser 1	Thr Ası 20	ı ser	GIN	Pne	тгр		Pne	Leu	Arg				
700			20					25							
701	(2) INFO	ם מאחדר	אר פרט	CEO	TD N	2.21									
702	(2) INFO	WMIIC	JN FOR	SEQ	ID M	J: ZI	•								
702	/; >	CEOU	ENCE CI	ים גם גנ	T C T T	em T C	٠.								
704	(+)	_	LENGTI												
705			TYPE:				cras								
706							1.								
707			TOPOL			_	īe								
708		(D)	TOPOL	JGI.	TIME	11									
709	1333	MOT E	ית כדווי	/DF.	nont:	: 40									
710	(11)	MOLEC	CULE T	.PE.	pept.	Lue									
711	11111	HVDOT	THETICA	AT • NI	^										
712	( +++)	mro	Inelica	TT . 14	•										
713															
714	(5511)	TMMET	DIATE S	יטוופכי	₽•										
715	( • ± ± )		CLONE	-		ia n	rodu	¬+ 17	6						
716		(5)	CLONE	. IIyu	Olys.	re b	Louu	- C V	0						
717															
718	(vi)	SEOUE	ENCE DI	CCPT	ו∩דתם	J. C1	ro ti	א א	. 21.						
719	(11)	22501		JONE			35 TI	, 110							
720	Glv	T.vg G	ly Ile	Agn	Tla	T.να	Ser	Ser	T.011	Sar	Glv	T.011	Thr	Va1	T.e.ii
721	1	_,	, , , , , ,	5 5		_,_	-	001	10	001	O <sub>1</sub>	DCu		15	DCu
722	_			•											
723	Lys														
724	-,,-														
725															
726	(2) INFO	RMATTC	N FOR	SEO	א מד	1.22	•								
727	(=, =:::0.			ULY.			•								
728	(1)	SEQUE	ENCE CI	IARAC	PERTS	STICS	3 •								
729	(-)	_	LENGT					=							
730			TYPE:					•							
731			STRANI												
732			TOPOLO												
733		(-,													
		MOLEC		יסקי	ONA A	aena	omic	١							
734	(ii)		JULE TY												
734 735	(ii)		OLE T		,	(90	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,								
735						( gen		•							
			HETICA			( gen									
735 736						, ge.i.									
735 736 737	(iii)	нүрот	THETICA	L: NO	o	, ye.i.									
735 736 737 738	(iii)	НҮРОТ	THETICA	L: NO	o E:										
735 736 737 738 739	(iii)	НҮРОТ	THETICA	L: NO	o E:										

793 794 795

#### Raw Sequence Listing

04/15/92 09:06:49

#### Patent Application US/07/659,408A

743	(X1) S	EQUENCE DESC	CRIPTION: SI	EQ ID NO:22:	•		
744 745	CATCCCCCCA	AGCATAAAGT	СТАВАСССТС	СССТСССТАА	<b>ТСАСТСАССТ</b>	<b>አ</b> አርሞጥልሮልጥጥ	60
746	onreceden	11001111111101	GIMMOCCIG	GGGTGCCTAIN	TONGTONGCT	AACTIACATI	
747 748	AATTGCGTTG	CGCTCACTGC	CCGCTTTCCA	GTCGGGAAAC	CTGTCGTGCC	AGCTGCATTA	120
749 750	ATGAATCGGC	CAACGCGCGG	GGAGAGGCGG	TTTGCGTATT	GGGCGCCAGG	GTGGTTTTTC	180
751 752	TTTTCACCAG	TGAGACGGGC	AACAGCTGAT	TGCCCTTCAC	CGCCTGGCCC	TGAGAGAGTT	240
753 754	GCAGCAAGCG	GTCCACGCTG	GTTTGCCCCA	CCACCGAAA	ATCCTGTTTG	ATGGTGGTTA	300
755	ACGGCGGGAT	ATAACATGAG	CTGTCTTCGG	TATCGTCGTA	TCCCACTACC	GAGATATCCG	360
756 757	CACCAACGCG	CAGCCCGGAC	TCGGTAATGG	CGCGCATTGC	GCCCAGCGCC	ATCTGATCGT	420
758 759 760	TGGCAACCAG	CATCGCAGTG	GGAACGATGC	CCTCATTCAG	CATTTGCATG	GTTTGTTGAA	480
761 762	AACCGGACAT	GGCACTCCAG	TCGCCTTCCC	GTTCCGCTAT	CGGCTGAATT	TGATTGCGAG	540
763 764	TGAGATATTT	ATGCCAGCCA	GCCAGACGCA	GACGCGCCGA	GACAGAACTT	AATGGGCCCG	600
765 766	CTAACAGCGC	GATTTGCTGG	TGACCCAATG	CGACCAGATG	CTCCACGCCC	AGTCGCGTAC	660
767 768	CGTCTTCATG	GGAGAAAATA	ATACTGTTGA	TGGGTGTCTG	GTCAGAGACA	TCAAGAAATA	720
769 770	ACGCCGGAAC	ATTAGTGCAG	GCAGCTTCCA	CAGCAATGGC	ATCCTGGTCA	TCCAGCGGAT	780
771 772	AGTTAATGAT	CAGCCCACTG	ACGCGTTGCG	CGAGAAGATT	GTGCACCGCC	GCTTTACAGG	840
773 774	CTTCGACGCC	GCTTCGTTCT	ACCATCGACA	CCACCACGCT	GGCACCCAGT	TGATCGGCGC	900
775 776	GAGATTTAAT	CGCCGCGACA	ATTTGCGACG	GCGCGTGCAG	GGCCAGACTG	GAGGTGGCAA	960
777 778	CGCCAATCAG	CAACGACTGT	TTGCCCGCCA	GTTGTTGTGC	CACGCGGTTG	GGAATGTAAT	1020
779 780	TCAGCTCCGC	CATCGCCGCT	TCCACTTTTT	CCCGCGTTTT	CGCAGAAACG	TGGCTGGCCT	1080
781 782	GGTTCACCAC	GCGGGAAACG	GTCTGATAAC	AGACACCGGC	ATACTCTGCG	ACATCGTATA	1140
783 784	ACGTTACTGG	TTTCACATTC	ACCACCCTGA	ATTGACTCTC	TTCCGGGCGC	TATCATGCCA	1200
785 786	TACCGCGAAA	GGTTTTGCGC	CATTCGATGG	TGTCCG			1236
787 788	(2) INFORM	ATION FOR SE	EQ ID NO:23:	:			
789	(i) S	EQUENCE CHAP	RACTERISTICS	3:			
790	, ,	(A) LENGTH:					
791		(B) TYPE: nu	-	~			
792		(C) STRANDED		le			
793		(D) TOPOLOGY		-			
		(-, -010000					

(ii) MOLECULE TYPE: DNA (genomic)

#### Raw Sequence Listing

04/15/92 09:06:56

796 797		(iii	) HY	РОТНІ	TIC	AT.: 1	NO										
798		<b>\</b>	,														
799																	
800		(vii	\ TMI	MEDIA	ATE S	SOUR	CE:										
801		, ,	•	B) CI				nt 4									
802			١,	J, U.		• • •	agc.	110 4									
803		/iv	\ FF:	ATURI													
804		127	•	A) N		zrv.	CDS										
805			•	B) L(	•			21	<u>د</u>								
806			•	•							·	II	.1			al + a	_
807			(1	J) U.								_		-	Bign	al T d	a
808					1	4.4 III	uman	gro	WCII !	IOTIN	one j	prec	arso.	<b>L</b>			
809																	
		1 2		<b></b>	30 DI	acon.	T D.M.T.	<b>.</b>	CEO :	TD 17		_					
810		(XI	) SE	QUENC	וע שנ	SOCK.	IPTIC	JN:	SEQ.	א עז	J: 23	:					
811	moo:	N C C C C	~~~ (	TO T O	-m-c-m-r	na ar			a	~~ » m·		m = m	ma	mam .			
812	TCG	AGCT	JAC :	IGAC	JTGT.	rg C	TTAT	ATTA	C AT	CGAT	AGCG	TAT	AATG	rGT (	GGAA.	TTGTGA	60
813			~~~		. ~ ~ ~ .			~~~~									
814	GCG	A'I'AA(	CAA !	rttc?	ACACA	AG T	l'TAA(	CTTT	A AG	AAGG	AGAT	ATA					115
815													ı	Met i	Ala '	rnr	
816														1			
817																	
818				ACT													163
819	GIĀ		Arg	Thr	ser	Leu		Leu	Ala	Phe	GIĀ		Leu	Cys	Leu	Pro	
820		5					10					15					
821																	
822				GAG													211
823	_	Leu	Gln	Glu	Gly		Ala	Phe	Pro	Thr		Pro	Leu	Ser	Arg		
824	20					25					30					35	
825																	
826				GCT													259
827	Phe	Asp	Asn	Ala		Leu	Arg	Ala	His		Leu	His	Gln	Leu		Phe	
828					40					45					50		
829																	
830				CAG													307
831	Asp	Thr	Tyr	Gln	Glu	Phe	Glu	Glu		Tyr	Ile	Pro	Lys		Gln	Lys	
832				55					60					65			
833																	
834				CTG	CA												321
835	Tyr	Ser	Phe														
836			70														
837																	
838																	
839	(2)	INF	DRMAT	rion	FOR	SEQ	ID I	NO: 2	4:								
840																	
841		4	(i) S	SEQUE													
842							: 70			cids							
843				(B)	TYI	PE: 8	amino	ac:	id								
844				(D)	TOE	POLO	3Y: 3	line	ar								
845				•													
846		( :	Li) N	OLEC	CULE	TYPE	E: pi	rote	in								
847		-	-				-										
848		()	(i) S	SEQUE	ENCE	DESC	CRIP	CION	: SEC	O ID	NO:2	24:					

#### Raw Sequence Listing

04/15/92 09:07:04

849																		
850	Vot	717	mh ∽	C1	50×	λ ~~	መb w	50×	T 011	T 011	T 0	710	Dho	C1	T 011	T 011		
851	met 1	nia	1111	GIY	5er 5	Arg	IIIL	Ser	Leu	10	Leu	HIG	Pne	GIY	15	Leu		
852	_				5					10					15			
853	Cvc	Len	Dro	Trn	LON	Gln	Glu	Clar	502	ת 1 ת	Dho	Dro	Thr	T10	Dro	Tou		
854	Cyb	Leu	PLO	20	Deu	GIII	Giu	GIY	25	nia	FILE	PLU	1111	30	PIO	Leu		
855				20					25					30				
856	50=	7~~	Tou	Dho	N an	Asn	λla	Vo+	T 011	7 ~~~	<b>71</b> ~	นเล	7~~	T 011	บเล	Cln		
857	ser	ALG	35	FIIE	nap	VPII	nia	40	Leu	ALG	MIG	ura	45	Leu	ura	GIII		
858			23					40					43					
859	T 011	712	Dho	200		Tyr	C1-	C1	Dho	C1	C1	<b>7.1</b> a	m	T10	Dwa	T		
860	Deu	50	FIIE	veb	1111	TYL	55	Giu	FIIC	GIU	GIU	60	-	TIE	PIO	гåа		
861		50					55					00						
862	Glu	Gln	T.v.a	Туг	Sar	Dhe												
863	65	0111	Lyo	-7-	Der	70												
864	05					70												
865	121	TNFC	יבאפר	TON	FOR	SEQ	ו חד	NO • 21	5•									
866	(2)	1111	J14111	11011	ron	סבע	10 .	10.2	٠.									
867		13	SEC	TIEN	רות כו	HARAG	ጎጥፎው ነ	TCTT	re.									
868		, -,		-		H: 74												
869			•	•		nuc		-										
870						DEDNI												
871			•			OGY:												
872			١.	, -														
873		(ii'	MOI	LECUI	LE TY	YPE:	DNA	(ge	nomi	2)								
874		,,	,					(90.		-,								
875		(iii'	HY	РОТНІ	ETTC	AL: 1	OV											
876		,	,															
877																		
878		(vii	IMI	MEDIA	ATE S	SOUR	CE:											
879						: Cla		deI :	frag	nent								
880			•	,														
881																		
882		(xi)	) SE	QUEN	CE DI	ESCRI	PTI	ON:	SEQ :	ID NO	0:25	:						
883				_					_									
884	CGA:	CAGC	TA :	TAAT	GTGT	GG A	ATTG	rgag(	C GG	ATAAC	CAAT	TTC	ACAC	AGT :	TTTT	CGCGAI	A	60
885																		
886	GAA	GAG	ATA :	<b>FACA</b>														74
887																		
888	(2)	INFO	DRMA!	rion	FOR	SEQ	ID I	NO:26	5:									
889	-																	
890		(i)	SE	QUEN	CE CI	HARAC	CTER	ISTI	cs:									
891			(2	A) LI	ENGTI	H: 19	90 ba	ase p	pairs	3								
892			(1	3) T	YPE:	nucl	leic	acio	Ĺ									
893			((	c) s:	<b>TRANI</b>	DEDNE	ESS:	doul	ole									
894			(I	) T(	OPOLO	OGY:	line	ear										
895																		
896		(ii)	MOI	LECUI	LE TY	PE:	DNA	(gei	nomi	=)								
897		•						-										
898	+	(iii)	HYI	POTHE	ETICA	AL: N	10											
899																		
900																		
901		(vii)	IMI	MEDIA	ATE S	SOURC	E:											

#### Raw Sequence Listing

04/15/92 09:07:11

902 903	(B) CLONE: Plasmid p373,2 fragment	
903		
905	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:	
906		
907 908	GATCTTCAAG CAGACCTACA GCAAGTTCGA CACAAACTCA CACAACGATG ACGCACTACT	60
909 910	CAAGAACTAC GGGCTGCTCT ACTGCTTCAG GAAGGACATG GACAAGGTCG AGACATTCCT	120
911	GCGCATCGTG CAGTGCCGCT CTGTGGAGGG CAGCTGTGGC TTCTAGTAAG GTACCCTGCC	180
912 913	CTACGTACCA	100
914	CTACGTACCA	190
915	(2) INFORMATION FOR SEQ ID NO:27:	
916	(2) INFORMATION FOR SEQ ID NO.27.	
917	(i) SEQUENCE CHARACTERISTICS:	
918	(A) LENGTH: 48 base pairs	
919	(B) TYPE: nucleic acid	
920	(C) STRANDEDNESS: single	
921	(D) TOPOLOGY: linear	
922	(0) 10102010 2011012	
923	(ii) MOLECULE TYPE: DNA (genomic)	
924	()	
925	(iii) HYPOTHETICAL: NO	
926	, ,	
927		
928	(vii) IMMEDIATE SOURCE:	
929	(B) CLONE: AccI-NdeI synthetic fragment	
930		
931		
932	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:	
933		
934 935	TATGTCTGCG GTAAAAGCAG CGCGCTACGG CAAGGACAAT GTTCGCGT	48
936 937	(2) INFORMATION FOR SEQ ID NO:28:	
938	(i) SEQUENCE CHARACTERISTICS:	
939	(A) LENGTH: 360 base pairs	
940	(B) TYPE: nucleic acid	
941	(C) STRANDEDNESS: single	
942	(D) TOPOLOGY: linear	
943		
944	(ii) MOLECULE TYPE: DNA (genomic)	
945		
946	(iii) HYPOTHETICAL: NO	
947		
948	(!!) TIMEDIAMO COMPON	
949	(vii) IMMEDIATE SOURCE:	
950	(B) CLONE: Plasmid pEMR469 fragment	
951 952		
952 953	(wi) CECHENCE DECCRIPATION, CEC TO NO. 20.	
954	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:	

#### Raw Sequence Listing

04/15/92 09:07:19

955 956	GGGACGCGTC TCCTCTGCCG GAACACCGGG CATCTCCAAC TTATAAGTTG GAGAAATAAG	60
957 958	AGAATTTCAG ATTGAGAGAA TGAAAAAAAA AAAAAAAAAA	120
959 960	AATGGGGTTC ACTTTTTGGT AAAGCTATAG CATGCCTATC ACATATAAAT AGAGTGCCAG	180
961 962	TAGCGACTTT TTTCACACTC GAGATACTCT TACTACTGCT CTCTTGTTGT TTTTATCACT	240
963 964	TCTTGTTTCT TCTTGGTAAA TAGAATATCA AGCTACAAAA AGCATACAAT CAACTATCAA	300
965 966	CTATTAACTA TATCGATACC ATATGGATCC GTCGACTCTA GAGGATCGTC GACTCTAGAG	360
967 968	(2) INFORMATION FOR SEQ ID NO:29:	
969		
970	(i) SEQUENCE CHARACTERISTICS:	
971	(A) LENGTH: 58 base pairs	
972	(B) TYPE: nucleic acid	
973	(C) STRANDEDNESS: double	
974	(D) TOPOLOGY: linear	
975		
976	(ii) MOLECULE TYPE: DNA (genomic)	
977		
978	(iii) HYPOTHETICAL: NO	
979		
980		
981	(vii) IMMEDIATE SOURCE:	
982	(B) CLONE: Fragment C	
983		
984		
985	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:	
986		
987	CGATATACAC AATGTCTGCT GTTAAGGCTG CTAGATACGG TAAGGACAAC GTTAGAGT	58
988		
989	(2) INFORMATION FOR SEQ ID NO:30:	
990		
991	(i) SEQUENCE CHARACTERISTICS:	
992	(A) LENGTH: 1013 base pairs	
993	(B) TYPE: nucleic acid	
994	(C) STRANDEDNESS: double	
995	(D) TOPOLOGY: linear	
996		
997	(ii) MOLECULE TYPE: DNA (genomic)	
998		
999	(iii) HYPOTHETICAL: NO	
1000		
1001		
1002	(vii) IMMEDIATE SOURCE:	
1003	(B) CLONE: Fragment D	
1004		
1005		
1006	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:	
1007		

1059

1060

#### Raw Sequence Listing

04/15/92 09:07:26

#### Patent Application US/07/659,408A

1008	CTACAAGGTT	CACAAGGACG	AGAAGACCGG	TGTCCAGACG	GTGTACGAGA	TGACCGTCTG	60
1009 1010	тстссттстс	GAGGGTGAGA	ттсасасстс	ттасассаас	CCCCACAACA	CCCTCATTCT	120
1011	1010011010	GAGGGTGAGA	TIGAGACCIC	TINCHCCANG	GCCGACAACA	GCGTCATIGI	120
1012	CGCAACCGAC	TCCATTAAGA	ACACCATTTA	CATCACCGCC	AAGCAGAACC	CCGTTACTCC	180
1013							
1014	TCCCGAGCTG	TTCGGCTCCA	TCCTGGGCAC	ACACTTCATT	GAGAAGTACA	ACCACATCCA	240
1015 1016	тессестеле	GTCAACATTG	<b>ጥርጥርርር እርርር</b>	СТССАССССС	<b>А</b> ТССАСАТТС	ACCCCAACCC	300
1017	IGCCGCICAC	GICAACAIIG	TOTOCCACCO	CIGGACCCGG	AIGGACAIIG	ACGGCAAGCC	300
1018	ACACCCTCAC	TCCTTCATCC	GCGACAGCGA	GGAGAAGCGG	AATGTGCAGG	TGGACGTGGT	360
1019							
1020	CGAGGGCAAG	GGCATCGATA	TCAAGTCGTC	TCTGTCCGGC	CTGACCGTGC	TGAAGAGCAC	420
1021 1022	GA N GMGGGA G	mmomococom	maamaaamaa	GG3.GE3.G3.GG	1 C1 CMM11 CC	1 C1 CCMCCC1	400
1022	CAACTCGCAG	TTCTGGGGCT	TCCTGCGTGA	CGAGTACACC	ACACTTAAGG	AGACCTGGGA	480
1024	CCGTATCCTG	AGCACCGACG	TCGATGCCAC	TTGGCAGTGG	AAGAATTTCA	GTGGACTCCA	540
1025							
1026	GGAGGTCCGC	TCGCACGTGC	CTAAGTTCGA	TGCTACCTGG	GCCACTGCTC	GCGAGGTCAC	600
1027	mama)	mmma.amaa					
1028 1029	TCTGAAGACT	TTTGCTGAAG	ATAACAGTGC	CAGCGTGCAG	GCCACTATGT	ACAAGATGGC	660
1030	AGAGCAAATC	CTGGCGCGCC	AGCAGCTGAT	CGAGACTGTC	GAGTACTCGT	TGCCTAACAA	720
1031		010000000		001101101101	00101001	100011110111	720
1032	GCACTATTTC	GAAATCGACC	TGAGCTGGCA	CAAGGGCCTC	CAAAACACCG	GCAAGAACGC	780
1033							
1034	CGAGGTCTTC	GCTCCTCAGT	CGGACCCCAA	CGGTCTGATC	AAGTGTACCG	TCGGCCGGTC	840
1035 1036	СТСТСТСААС	ጥሮሞልልልጥጥርጥ	AAACCAACAT	CATTCTCACC	<b>ጥጥርርርር አርጥጥ</b>	TCCAAGGCAA	900
1037	CICICIOAAG	ICIAAAIIGI	AMCCANCAI	GATICICACG	TICCGGAGII	TCCAAGGCAA	300
1038	ACTGTATATA	GTCTGGGATA	GGGTATAGCA	TTCATTCACT	TGTTTTTTAC	TTCCAAAAAA	960
1039							
1040	AAAAAAAAA	AAAAAAAAA	AAAAAAAAA	AAAAAAAAA	AAAAAAGGGC	CCG	1013
1041 1042	(2) INFORM	AMTON BOD CE	FO TD NO. 21.				
1042	(2) INFORM	ATION FOR SE	SQ ID NO:31:				
1044	(i) SE	EQUENCE CHAP	RACTERISTICS	S:			
1045	` '	(A) LENGTH:	207 base pa	airs			
1046	(	(B) TYPE: nu	cleic acid				
1047		(C) STRANDEI		le			
1048	(	(D) TOPOLOGY	: linear				
1049							
1050	(ii) MC	DLECULE TYPE	E: DNA (gend	omic)			
1051							
1052	(111) HA	POTHETICAL:	טא :				
1053 1054							
1054	/wiil Th	MEDIATE SOU	IDCF•				
1055		(B) CLONE: S		AI.7 frament			
1057	,	(D) CHORE. S	Augustie CTC G	ar, rradment	•		
1058							
1000							

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

#### Raw Sequence Listing

04/15/92 09:07:33

1061 1062	CGCGTCTATA CTTCGGAGCA CTGTTGAGCG AAGGCTCATT AGATATATTT TCTGTCATTT	60
1063 1064	TCCTTAACCC AAAAATAAGG GAGAGGGTCC AAAAAGCGCT CGGACAACTG TTGACCGTGA	120
1065 1066	TCCGAAGGAC TGGCTATACA GTGTTCACAA AATAGCCAAG CTGAAAATAA TGTGTAGCCT	180
1067 1068	TTAGCTATGT TCAGTTAGTT TGGCATG	207
1069 1070	(2) INFORMATION FOR SEQ ID NO:32:	
1071	(i) SEQUENCE CHARACTERISTICS:	
1072	(A) LENGTH: 23 base pairs	
1072	(B) TYPE: nucleic acid	
1073	· ·	
1074	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
1076	(11) VOLDOUT D. BUDD. BUD. (consents)	
1077	(ii) MOLECULE TYPE: DNA (genomic)	
1078	4111	
1079	(iii) HYPOTHETICAL: NO	
1080		
1081		
1082	(vii) IMMEDIATE SOURCE:	
1083	(B) CLONE: Modified XbaI-MluI adapter	
1084		
1085		
1086	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:	
1087		
1088	CTAGGCTAGC GGGCCCGCAT GCA	23
1089		
1090	(2) INFORMATION FOR SEQ ID NO:33:	
1091		
1092	(i) SEQUENCE CHARACTERISTICS:	
1093	(A) LENGTH: 422 base pairs	
1094	(B) TYPE: nucleic acid	
1095	(C) STRANDEDNESS: single	
1096	(D) TOPOLOGY: linear	
1097		
1098	(ii) MOLECULE TYPE: DNA (genomic)	
1099		
1100	(iii) HYPOTHETICAL: NO	
1101	• •	
1102		
1103	(vii) IMMEDIATE SOURCE:	
1104	(B) CLONE: Plasmid pSE1 "site binding to HindIII"	
1105	fragment	
1106	· <b>-</b> · · · · · · · · · · · · · · · · · · ·	
1107		
1108	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:	
1109	//	
1110	100000000000000000000000000000000000000	60
1111	AGCTGGCTCG CATCTCTCCT TCACGCGCCC GCCGCCCTAC CTGAGGCCGC CATCCACGCC	
	AGCTGGCTCG CATCTCTCT TCACGCGCCC GCCGCCCTAC CTGAGGCCGC CATCCACGCC	
1111 1112 1113	GGTGAGTCGC GTTCTGCCGC CTCCCGCCTG TGGTGCCTCC TGAACTGCGT CCGCCGTCTA	120

#### Raw Sequence Listing

04/15/92 09:07:41

1114 1115	GGTAGGCTCC AAGGGAGCCG GACAAAGGCC CGGTCTCGAC CTGAGCTCTA AACTTACCTA	180
1116 1117	GACTCAGCCG GCTCTCCACG CTTTGCCTGA CCCTGCTTGC TCAACTCTAC GTCTTTGTTT	240
1118 1119	CGTTTTCTGT TCTGCGCCGT TACAACTTCA AGGTATGCGC TGGGACCTGG CAGGCGGCAT	300
1120 1121	CTGGGACCCC TAGGAAGGC TTGGGGGTCC TCGTGCCCAA GGCAGGGAAC ATAGTGGTCC	360
1122 1123	CAGGAAGGGG AGCAGAGGCA TCAGGGTGTC CACTTTGTCT CCGCAGCTCC TGAGCCTGCA	420
1124	GA	422
1125		
1126	(2) INFORMATION FOR SEQ ID NO:34:	
1127		
1128	(i) SEQUENCE CHARACTERISTICS:	
1129	(A) LENGTH: 77 base pairs	
1130	(B) TYPE: nucleic acid	
1131	(C) STRANDEDNESS: double	
1132	(D) TOPOLOGY: linear	
1133		
1134	(ii) MOLECULE TYPE: DNA (genomic)	
1135		
1136	(iii) HYPOTHETICAL: NO	
1137		
1138		
1139	(vii) IMMEDIATE SOURCE:	
1140	(B) CLONE: Synthetic HindIII-"site binding to BamHI"	
1141	fragment	
1142	$\cdot$	
1143		
1144	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:	
1145	100mmcmoo1	
1146 1147	AGCTTGTCGA CTAATACGAC TCACTATAGG GCGCCGCGG GCCCCTGCAG GAATTCGGAT	60
1147	000000000 3 0mo3 cm	
1149	CCCCGGGTG ACTGACT	77
1150	(2) INFORMATION FOR SEQ ID NO:35:	
1151	(2) INFORMATION FOR SEQ ID NO:55:	
1152	(i) SEQUENCE CHARACTERISTICS:	
1153	(A) LENGTH: 61 base pairs	
1154	(B) TYPE: nucleic acid	
1155	(C) STRANDEDNESS: double	
1156	(D) TOPOLOGY: linear	
1157	(5) 10101001. 111001	
1158	(ii) MOLECULE TYPE: DNA (genomic)	
1159	(-5, 55-25-25-25-5-5-5-5-5-5-5-5-5-5-5-5-5-5	
1160	(iii) HYPOTHETICAL: NO	
1161	( ) ,	
1162		
1163	(vii) IMMEDIATE SOURCE:	
1164	(B) CLONE: Synthetic HindIII-AccI fragment	
1165		
1166		

#### Raw Sequence Listing

04/15/92 09:07:48

1167	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:	
1168		-
1169	AGCTTGCCGC CACTATGTCC GCAGTAAAAG CAGCCCGCTA CGGCAAGGAC AATGTCCGCG	60
1170 1171	T	61
1172	•	0.1
1172	(2) INFORMATION FOR SEQ ID NO:36:	
1174	(2) INFORMATION FOR SEQ ID NO:30:	
1175	(i) SEQUENCE CHARACTERISTICS:	
1176	(A) LENGTH: 920 base pairs	
1177	(B) TYPE: nucleic acid	
1178	(C) STRANDEDNESS: single	
1179	(D) TOPOLOGY: linear	
1180	• •	
1181	(ii) MOLECULE TYPE: DNA (genomic)	
1182		
1183	(iii) HYPOTHETICAL: NO	
1184		
1185		
1186	(vii) IMMEDIATE SOURCE:	
1187	(B) CLONE: HindIII-SnaBI fragment	
1188		
1189		
1190	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:	
1191		
1192	AGCTTGCCGC CACTATGTCC GCAGTAAAAG CAGCCCGCTA CGGCAAGGAC AATGTCCGCG	60
1193		
1194 1195	TCTACAAGGT TCACAAGGAC GAGAAGACCG GTGTCCAGAC GGTGTACGAG ATGACCGTCT	120
1195		100
1196	GTGTGCTTCT GGAGGGTGAG ATTGAGACCT CTTACACCAA GGCCGACAAC AGCGTCATTG	180
1197	TCGCAACCGA CTCCATTAAG AACACCATTT ACATCACCGC CAAGCAGAAC CCCGTTACTC	240
1199	TOGERACEON CICCATTANG ANCACCATTI ACATCACCGC CAAGCAGAAC CCCGTTACTC	240
1200	CTCCCGAGCT GTTCGGCTCC ATCCTGGGCA CACACTTCAT TGAGAAGTAC AACCACATCC	300
1201	Transfer direction and transfer in the analysis and the contraction	300
1202	ATGCCGCTCA CGTCAACATT GTCTGCCACC GCTGGACCCG GATGGACATT GACGGCAAGC	360
1203		550
1204	CACACCCTCA CTCCTTCATC CGCGACAGCG AGGAGAAGCG GAATGTGCAG GTGGACGTGG	420
1205		
1206	TCGAGGGCAA GGGCATCGAT ATCAAGTCGT CTCTGTCCGG CCTGACCGTG CTGAAGAGCA	480
1207		
1208	CCAACTCGCA GTTCTGGGGC TTCCTGCGTG ACGAGTACAC CACACTTAAG GAGACCTGGG	540
1209		
1210	ACCGTATCCT GAGCACCGAC GTCGATGCCA CTTGGCAGTG GAAGAATTTC AGTGGACTCC	600
1211		
1212	AGGAGGTCCG CTCGCACGTG CCTAAGTTCG ATGCTACCTG GGCCACTGCT CGCGAGGTCA	660
1213		
1214	CTCTGAAGAC TTTTGCTGAA GATAACAGTG CCAGCGTGCA GGCCACTATG TACAAGATGG	720
1215		
1216	CAGAGCAAAT CCTGGCGCGC CAGCAGCTGA TCGAGACTGT CGAGTACTCG TTGCCTAACA	780
1217		
	AGCACTATTT CGAAATCGAC CTGAGCTGGC ACAAGGGCCT CCAAAACACC GGCAAGAACG	840
1219		

Page: 24 Raw Sequence Listing 04/15/92 09:07:55

1220	CCGAGGTCTT	CGCTCCTCAG	TCGGACCCCA	ACGGTCTGAT	CAAGTGTACC	GTCGGCCGGT	900
1221							
1222	CCTCTCTGAA	СТСТАВАТТС					920

PAGE: 1

# SEQUENCE VERIFICATION REPORT PATENT APPLICATION US/07/659,408A

DATE: 04/15/92 TIME: 09:07:56

LINE ERROR

#### ORIGINAL TEXT

38	Wrong application Serial Numb	er (A) APPLICATIO	N NUMBER: US 07/659,408
221	Response Exceeds Line Limitat	ions prokaryotes	_
271	Response Exceeds Line Limitat	ions eukaryotes	
323	Response Exceeds Line Limitat	ions animal cells	(2)
345	Response Exceeds Line Limitat	ions cells	
1105	Response Exceeds Line Limitat	ions fragment	
1141	Response Exceeds Line Limitat	ions fragment	

PAGE: 1

### SEQUENCE MISSING ITEM REPORT PATENT APPLICATION US/07/659,408A

DATE: 04/15/92 TIME: 09:07:56

MANDATORY IDENTIFIER THAT WAS NOT FOUND

PRIOR APPLICATION DATA APPLICATION NUMBER FILING DATE

PAGE: 1

# SEQUENCE CORRECTION REPORT PATENT APPLICATION US/07/659,408A

CORRECTED TEXT

DATE: 04/15/92 TIME: 09:07:56

LINE ORIGINAL TEXT